In our efforts for cloning novel I2-superfamily conotoxins using the signal peptide sequence, we identified a novel conotoxin Lt12.4 from Conus litteratus. This gene has a framework XII (-C-C-C-CC-C-C-), which is distinct from the cysteine pattern I2-superfamily conotoxin (-C-CC-CC-C-C-C-). Subsequently, we found the signal peptide sequence of Lt12.4 by 5'-RACE. Using this new sequence, we identified another five novel conotoxins with this cysteine pattern from four Conus species (Conus eburneus, Conus imperialis, Conus marmoreus, and C. litteratus). These novel conotoxins have the same cysteine pattern as the reported Gla-TxX and Gla-MII, and may contain Gla residues. Furthermore, they have the highly conserved signal peptide and hypervariable mature peptide sequences, and widely exist in Conus species. Therefore, it could be defined as a new superfamily of E-conotoxins.

**Keywords** framework XII; conopeptides; cDNA clone; Conus; E-conotoxins

In our efforts for cloning novel I2-superfamily conotoxins using the signal peptide sequence, we identified a novel conotoxin Lt12.4 from Conus litteratus. This gene has a framework XII (-C-C-C-CC-C-C-), which is distinct from the cysteine pattern I2-superfamily conotoxin (-C-CC-CC-C-C-C-). Subsequently, we found the signal peptide sequence of Lt12.4 by 5'-RACE. Using this new sequence, we identified another five novel conotoxins with this cysteine pattern from four Conus species (Conus eburneus, Conus imperialis, Conus marmoreus, and C. litteratus). These novel conotoxins have the same cysteine pattern as the reported Gla-TxX and Gla-MII, and may contain Gla residues. Furthermore, they have the highly conserved signal peptide and hypervariable mature peptide sequences, and widely exist in Conus species. Therefore, it could be defined as a new superfamily of E-conotoxins.

**Keywords** framework XII; conopeptides; cDNA clone; Conus; E-conotoxins

In our efforts for cloning novel I2-superfamily conotoxins using the signal peptide sequence, we identified a novel conotoxin Lt12.4 from Conus litteratus. This gene has a framework XII (-C-C-C-CC-C-C-), which is distinct from the cysteine pattern I2-superfamily conotoxin (-C-CC-CC-C-C-C-). Subsequently, we found the signal peptide sequence of Lt12.4 by 5'-RACE. Using this new sequence, we identified another five novel conotoxins with this cysteine pattern from four Conus species (Conus eburneus, Conus imperialis, Conus marmoreus, and C. litteratus). These novel conotoxins have the same cysteine pattern as the reported Gla-TxX and Gla-MII, and may contain Gla residues. Furthermore, they have the highly conserved signal peptide and hypervariable mature peptide sequences, and widely exist in Conus species. Therefore, it could be defined as a new superfamily of E-conotoxins.

**Keywords** framework XII; conopeptides; cDNA clone; Conus; E-conotoxins

In our efforts for cloning novel I2-superfamily conotoxins using the signal peptide sequence, we identified a novel conotoxin Lt12.4 from Conus litteratus. This gene has a framework XII (-C-C-C-CC-C-C-), which is distinct from the cysteine pattern I2-superfamily conotoxin (-C-CC-CC-C-C-C-). Subsequently, we found the signal peptide sequence of Lt12.4 by 5'-RACE. Using this new sequence, we identified another five novel conotoxins with this cysteine pattern from four Conus species (Conus eburneus, Conus imperialis, Conus marmoreus, and C. litteratus). These novel conotoxins have the same cysteine pattern as the reported Gla-TxX and Gla-MII, and may contain Gla residues. Furthermore, they have the highly conserved signal peptide and hypervariable mature peptide sequences, and widely exist in Conus species. Therefore, it could be defined as a new superfamily of E-conotoxins.
Materials and Methods

Reagents
T4 DNA ligase, X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside), and IPTG (isopropyl-thio-β-D-galactoside) were purchased from SABC (Sino-American Biotechnology Company, Luoyang, China). The pGEM-T Easy vector system and gel purification kit Wizard DNA Clean-Up system for PCR product were obtained from Promega (Madison, USA). PrimeSTAR HS DNA Polymerase, 3'-Full RACE Core Set Ver.2.0 kit, and 5'-Full RACE kit were from TaKaRa (Dalian, China). The RNeasy Mini Kit was purchased from QIAGEN (Hilden, Germany). The other reagents were of analytical grade from SABC.

Isolation of total RNA
Specimens of four Conus species (C. eburneus, C. imperialis, C. marmoreus, and C. litteratus) were collected from an area around the Hainan Island in the South China Sea. The venom duct was dissected from the snails and immediately frozen in liquid nitrogen. Total RNA was extracted and purified from 100 mg of homogenized venom duct tissue from each Conus species by the RNeasy Mini Kit (QIAGEN).

3'-RACE
One microgram of the total venom duct RNA from the four Conus snails was reversely transcribed into cDNA by reverse transcriptase M-MLV with a 3'-RACE adaptor (5'-TACCGTGTTCCACTAGTGATTTTTTTTTTTTTTTTTT-3'). The synthesized cDNA was used as the template for the 3'-RACE reaction. The following primers were used: the forward primer F1 (5'-ATGATGTTTTCGATCTAGTG-3') was designed based on the N terminus of the signal sequence of I2-superfamily conotoxins, which was used to find the novel conotoxin Lt12.4 [25]; the forward outer primer P1 (5'-ATGTTGCGGTATACGTCAG-3') and the inner primer P2 (5'-CTTCTGTCTCTCCATCT-3') were designed based on the N-terminal Lt12.4 signal sequence, which were used to identify other novel conotoxins with framework XII. The reverse 3'-RACE outer primer R1 (5'-TACCGTGTTCCACTAGTGATT-3') was a shorter version of the 3'-RACE adaptor primer without the poly (dT) tail as well as reverse 3'-RACE inner primer R2 (5'-CGCGGATCCTCCACTAGTGATTTCATAG-3'). Briefly, PCR amplification was carried out with 30 cycles at melting temperature of 94°C for 30 s, annealing temperature of 55°C for 30 s, and extension temperature of 72°C for 45 s. Then, PCR products were analyzed by the gel electrophoresis, and the purified product of the expected size was added to adenine and ligated into the T-tailed plasmid vector pGEM-T (Promega). Subsequently, the ligation products were transformed into competent cells of Escherichia coli DH5α. Transformed colonies were screened by white–blue identification for sequence analysis. Plasmids containing inserts of ~300–700 bp in size were sequenced with the dideoxy chain termination method using the ABI Model 3730 automated sequencer (Applied Biosystems, Foster City, USA). The predicted protein sequences were analyzed with the software Seqtools (http://www.bio-soft.net/sms).

5'-RACE
The gene-specific primers D1 (5'-TAGGCTATGCAGGTGTTGTGCTG-3') and D2 (5'-CGGGTACATGCTGACAGCTA-3') were designed and synthesized based on the partial cDNA sequence from the 3' terminal of Lt12.4 and Mr12.5 obtained by 3'-RACE. The 5'-RACE outer primer (5'-CATGGCTACATGCTGACAGCTA-3') was provided in the 5'-RACE kit. The 5'-RACE was carried out using the TaKaRa 5'-RACE kit. Briefly, the 5'-cap of mRNA was isolated by the tobacco acid pyrophosphatase, and it was reversely transcribed into cDNA by reverse transcriptase M-MLV. The cDNA was used as the template for gene-specific PCR amplification with the gene-specific primer D1, D2, and 5'-RACE outer primer. Two prominent PCR bands of 600 and 397 bp were separated and purified. The product was ligated into a pGEM-T Easy vector (Promega) for sequencing. The signal peptide sequences ofLt12.4 and Mr12.5 were predicted by using the SignalP 3.0 server (http://www.cbs.dtu.dk/services/SignalP) [27].

Identification the framework XII conotoxins in other Conus species
Total cDNA of other Conus species was used as the template to clone other conotoxins of framework XII. The forward outer primer P1 and the inner primer P2 were designed based on the cDNA sequence of Lt12.4 signal peptide, and they were paired with the 3'-RACE outer primer and the 3'-RACE inner primer, respectively, for nested-PCR amplification. The PCR products were cloned into pGEM-T-Easy vector for sequencing. Finally, the novel cDNA sequences were searched in the GenBank Nucleotide Sequence Database.

Peptide sequence alignment and phylogenetic tree analysis
Database searches were performed with the BLASTN algorithm (http://www.ncbi.nlm.nih.gov/BLAST/) [28]. The predicted protein sequences were analyzed with the software Seqtools (http://www.bio-soft.net/sms) and the signal peptide sequences of the novel conotoxin precursors were predicted online with SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP/). Multiple alignments of protein...
sequences were obtained from CLUSTALX1.83. The phylogenetic analysis of new E-conotoxin sequences and the reported typical sequences with eight cysteines in the NCBI database was performed with the neighbor-joining (NJ) method by MEGA 4.0. Bootstrap values were estimated from 1000 replicates.

Results

Identification of Lt12.4 cDNA clones encoding conopeptide precursors with framework XII from C. litteratus

In our previous study, we identified a new conotoxin Lt12.4 (GenBank accession number GQ479944) using the conserved signal peptide sequence of the I2-superfamily [25], which has the framework XII. Figure 1 shows the full-length Lt12.4 cDNA cloned by the 3'- and 5'-RACE. The signal peptide sequence, post peptide, and mature toxin region of Lt12.4 consisted of 25, 14, and 46 amino acids, respectively. The mature toxin region was followed by post peptide and a closing 3'-UTR, which is similar to that of I2-superfamily. The signal peptide sequence of Lt12.4 is different from the I2-superfamily (only 52% of sequence identity) (Fig. 2), but it has a highly conserved sequence with Gla-TxX (84.6% of identity) [14]. Therefore, this led us to clone a new superfamily conotoxin from four Conus species with this signal peptide sequence of Lt12.4.

Identification of cDNA clones encoding framework XII conotoxins from four Conus species

We identified the cDNA sequences from the four Conus species (C. eburneus, C. imperialis, C. litteratus, and C. marmoreus) using 3'-RACE. We completely sequenced 68 clones in total and the inferred corresponding protein sequences were deduced with Seqtools software. In order to avoid experimental errors, we selected 16–18 clones for each Conus species, which were obtained from three different batches of PCR ligation products. We identified five new conotoxin sequences with the framework XII (-C-C-C-C-C-C-C-) from the four Conus species (one from C. eburneus, one from C. litteratus, two from C. marmoreus, one from C. imperialis, and C. eburneus) by searching the NCBI database. We named the corresponding mature peptides as conotoxin Im12.10, Lt12.9, Eb12.4, Mr12.5, Mr12.8, and Eb12.10. These cDNA sequences have been submitted to GenBank and the corresponding accession numbers are as follows: GQ228836 (Im12.10), GQ228838 (Lt12.9), GQ228839 (Eb12.4), GQ228841 (Mr12.5), and GQ228842 (Mr12.8). Figure 2 shows the predicted amino acid sequences. Surprisingly, our results showed that some precursor sequences were found to be identical in different Conus species. For example, Im12.10 was identical with Eb12.10, and the mature peptide of Eb12.4 from C. eburneus is also the same as Mr12.5 and the reported conotoxin Gla-MII from C. marmoreus [26].

Characterization of six new conotoxins of framework XII from four Conus species

Among the above six mature peptides, all the sequences have the same cysteine pattern and the difference among them is the number of amino acids between the second and the third cysteine. Interestingly, the mature peptide sequences exhibit both conserved and hypervariable characters between Conus species or within the same Conus species. For example, Lt12.4 and Im12.10 are from different Conus species and have significantly different

![Figure 1](https://example.com/figure1.png)
amino acid residues within the intercysteine loop. Lt12.4 and Lt12.9 are also different in residues, although they come from the same Conus species. On the other hand, we found the same conotoxins from different Conus species, such as Mr12.5 and Eb12.4.

**Discussion**

In the present study, we identified a new conotoxin Lt12.4 with the framework XII (C-C-C-C-CC-C-C-) during the cloning of the I2-superfamily conotoxins. By 5'-RACE, we cloned the signal peptide sequence of Lt12.4, which was different from the I2-superfamily conotoxins. However, it shows a more conserved sequence with Gla-TxX from C. textile. Furthermore, we identified the five other novel conotoxins with framework XII by using the signal peptide sequence of Lt12.4 (Fig. 2). The six novel members (Eb12.4, Lt12.9, Lt12.4, Mr12.5, Mr12.8, and Im12.10) along with the two reported conotoxins show more conserved signal peptide sequences, which are distinct from those sequences in previously reported superfamilies. Therefore, they can be defined as a novel class of E-superfamily conotoxins.

We cloned the novel Lt12.4 with this framework XII using the I2-superfamily signal peptide sequence. This success is mainly because of the moderate homology (52%) (Fig. 2) of the Lt12.4 signal peptide sequence with the I2 superfamily peptide as well as the similarity in the cloning primer F1, where the 15 bases of primer F1 (5'-ATGATGGTTTCTGAATGCATGTC-3') are identical to the signal peptide sequence of Lt12.4 (Fig. 1). Such cases have been observed in the clones of two O-superfamily conotoxins when the I3-superfamily signal sequence is used [13].

Similar to other superfamily conotoxins, there are hyper-variable mature peptide sequences in the novel six conotoxins. For example, Lt12.4 and Im12.10 are obviously different from other conotoxins (Fig. 2). A high conservation in precursor sequences (including signal peptide, mature peptide, and post-peptide) is in novel conotoxins. For example, Im12.10 exists in C. imperialis and C. eburneus (namely Eb12.10). Furthermore, the signal peptide sequence of Eb12.4 is different from Mr12.5 (reported as Gla-MII), but their mature peptides are the same. This phenomenon is rare, although a few similar cases have been reported [25,29].

To better understand these similarities of sequences and the evolutionary relationships among E-conopeptides and the other eight cysteine pattern conotoxins, we reconstructed a NJ tree computed from some typical eight cysteine conotoxins and E-superfamily precursors (Fig. 3). Among the six novel conotoxins, Lt12.9 and Im12.10 show a highly conserved sequence (the pairwise P-distance value is 0.36) and are clustered into one clade, the same evolutionary trace for Mr12.8, Eb12.4, and Mr12.5 (Gla-MII) (P-distance value is 0.00). In contrast, Lt12.4 belongs to distinct clade in the tree because of the different

---

**Figure 2** Precursor amino acid sequences

Alignment of amino acid sequences of six novel precursors from four Conus species and two reported precursors Gla-TxX and ViTx from C. textile and C. virgo, respectively. Clones from cDNA prepared from venom ducts of C. eburneus (Eb), C. imperialis (Im), C. litteratus (Lt), and C. marmerous (Mr) were obtained using PCR and RACE. The signal peptides were shaded and mature peptides were boxed. The predicted γ-carboxylation recognition sites were highlighted in white text on black background.

**Figure 3** Evolutionary relationships of eight cysteine peptides

The CLUSTALW1.83 multiple sequence alignment was used as an input for MEGA4. The probable evolutionary history was inferred using the NJ method. The bootstrap consensus tree from 1000 replicates is taken to represent the evolutionary history of the sequences analyzed. The percentages of replicate trees in which the associated sequences clustered together in the bootstrap test are shown next to the branches. Underscoring denotes the novel conopeptides identified in this study.
amino acid residues and numbers of amino acids occurring between the cysteine residues (the pairwise $P$-distance values between Lt12.4 and other five novel conotoxins range from 0.51 to 0.80). The different primary sequences may indicate different structure and biological functions. Furthermore, the other eight cysteine framework conotoxins show different cysteine patterns and amino acids; therefore, they are clustered into different clades. Interestingly, Ca11.1 (framework XVI) [17] and Ca11.1 (I$_3$-superfamily) [13] have distinct cysteine patterns, but they are clustered into the same clades (the pairwise $P$-distance value is 0.80). A similar case occurs in Lt115a (framework XV) [16] and ViTx (I$_2$-superfamily) [22].

Gla-TxX contains $\gamma$-carboxylated Glu residues in the mature peptide sequence and possesses a possible $\gamma$-carboxylation recognition site ($\gamma$-CRS, Lys/Arg-X-X-J-X-X-X-Lys/Arg, where $J$ is a hydrophobic residue, usually Leu, and $X$ is any amino acid) in the post-peptide sequence [14]. Eb12.4, Lt12.9, Mr12.5, Mr12.8, and Im12.10 contain the same Glu array and a similar conserved motif for $\gamma$-carboxylation recognition, and they probably have the same modification tendency with Gla-TxX, in which five Glu residues at loci 5, 9, 12, 15, and 16 of the mature peptide are post-translationally modified. In contrast, Lt12.4 has Glu residues at positions 4, 9, 21, suggesting that the $\gamma$-carboxylation site may be different from that of Gla-TxX. Previous studies have reported that the Gla-containing conotoxins can interact with Ca$^{2+}$ ions and inhibit calcium channel and NMDA receptor function [30,31]. So far, only two peptides with framework XII have been functionally characterized: one is an inhibitor of the porcine sodium/potassium ATPase [32], the other is a human elastase-specific inhibitor [33]. Thus, although the molecular target and the biological role remain to be determined, the framework XII conotoxin we identified here may be highly valuable.

Acknowledgements

We would like to thank Prof. Mary Prorok from the Department of Chemistry and Biochemistry at the University of Notre Dame, USA, for her helpful discussions.

Funding

This study was supported by grants from the High Technology Program of Oceans in China (2006AA09Z 404), the National Basic Research Program of China (2010CB529802), and the National Natural Science Foundation of China (30672446, 90713028).

References

unusual posttranslational modifications reduces presynaptic Ca\(^{2+}\) influx. 